

On the scaling of probability density functions with apparent power-law exponents less than unity

Kim Christensen^{1,2}, Nadia Farid², Gunnar Pruessner^{2,3a}, and Matthew Stapleton²

¹ Institute for Mathematical Sciences, Imperial College London, 53 Prince's Gate, London SW7 2PG, United Kingdom

² Blackett Laboratory, Imperial College London, Prince Consort Road, London SW7 2AZ, United Kingdom

³ Mathematics Institute, University of Warwick, Gibbet Hill Road, Coventry CV4 7AL, United Kingdom

February 3, 2008

Abstract. We derive general properties of the finite-size scaling of probability density functions and show that when the apparent exponent $\tilde{\tau}$ of a probability density is less than 1, the associated finite-size scaling ansatz has a scaling exponent τ equal to 1, provided that the fraction of events in the universal scaling part of the probability density function is non-vanishing in the thermodynamic limit. We find the general result that $\tau \geq 1$ and $\tau \geq \tilde{\tau}$. Moreover, we show that if the scaling function $\mathcal{G}(x)$ approaches a non-zero constant for small arguments, $\lim_{x \rightarrow 0} \mathcal{G}(x) > 0$, then $\tau = \tilde{\tau}$. However, if the scaling function vanishes for small arguments, $\lim_{x \rightarrow 0} \mathcal{G}(x) = 0$, then $\tau = 1$, again assuming a non-vanishing fraction of universal events. Finally, we apply the formalism developed to examples from the literature, including some where misunderstandings of the theory of scaling have led to erroneous conclusions.

Key words. Scaling – power laws – power law exponent less than unity – critical phenomena

PACS. 89.75.Da Systems obeying scaling laws – 89.75.-k Complex systems – 05.65.+b Self-organized systems – 89.75.Hc Networks and genealogical trees – 05.70.Jk Critical point phenomena

1 Introduction

Power-law probability densities are pervasive in the literature on critical and scale-invariant systems [1–3]. The supposedly sole scaling parameter in a finite system at the critical point is the *upper cutoff* s_c , diverging in the thermodynamic limit, $L \rightarrow \infty$, and the probability density function (PDF) $P(s; s_c)$ for an *observable* (an event size or the order parameter) $s \in [0, \infty[$ obeys simple *finite-size scaling* (FSS), that is,

$$P(s; s_c) = as^{-\tau} \mathcal{G}(s/s_c) \quad \text{for } s, s_c \gg s_0, \quad (1)$$

where s_0 is a constant *lower cutoff* of the PDF. Equation (1) is valid only in the region $s \gg s_0$ where the corrections to scaling are negligible [4]. The dimensionful parameter a is a so-called non-universal *metric factor* [5], τ is a universal (critical) *scaling exponent*, \mathcal{G} is the universal *scaling function* that decays sufficiently fast for $s \gg s_c$ to ensure that, as one would expect, all moments of the PDF are finite for finite system size, $L < \infty$. The upper cutoff s_c is the *characteristic size* of s in a finite system. Usually $s_c = bL^D$ to leading order, where b is another non-universal *metric factor* and D is the universal

spatial dimensionality of the observable s . Again, in general, there are sub-leading orders, that is, corrections to scaling of the form $s_c = bL^D (1 + b_1 L^{-\omega_1} + b_2 L^{-\omega_2} + \dots)$ with $0 < \omega_1 < \omega_2 < \dots$, which can be safely ignored in the analysis of the asymptotes. The scaling exponent τ is *uniquely defined* by Eq. (1) and is associated with the power-law decay of the distinctive onset of the rapid decay $P(s_c; s_c) = a\mathcal{G}(1)s_c^{-\tau}$ as a function of s_c . If data are consistent with the FSS ansatz (1), we can perform a *data collapse*: By plotting the transformed PDF $s^\tau P(s; s_c)$ vs. the rescaled observable s/s_c , all data for $s, s_c \gg s_0$ collapse onto the same curve representing the scaling function \mathcal{G} , see Figure 1. However, due to the presence of the metric factor a , the scaling function \mathcal{G} is determined by Eq. (1) only up to a prefactor. This ambiguity can be resolved, for example by fixing $\mathcal{G}(1) = 1$ [6].

The upper and lower cutoffs define a *scaling region* $s_0 \ll s \ll s_c$ where the PDF shows essentially power-law behaviour, rather informally

$$P(s; s_c) \propto s^{-\tilde{\tau}} \quad \text{for } s_0 \ll s \ll s_c, \quad (2)$$

which means that $s^{-\tilde{\tau}}$ is the leading order of $P(s; s_c)$ in the region where $s \gg s_0$ coexists with $s \ll s_c$. The *apparent exponent* $\tilde{\tau}$ is the slope of a straight-line fit to the PDF data in the scaling region $s_0 \ll s \ll s_c$ when plotting $\log P(s; s_c)$ vs. $\log s$, see Figure 1(a).

^a Present address: Department of Mathematics, Imperial College London, 180 Queen's Gate, London SW7 2AZ, United Kingdom

Let us reflect upon the FSS ansatz (1) and the informal Eq. (2). Equation (2) is informal for various reasons. First, the “ \propto ” suggests a proportionality without specifying what parameters a possible prefactor would be allowed to depend on. Second, the upper cutoff appears in the condition $s \ll s_c$, but s_c does not enter Eq. (2) explicitly as it does in Eq. (1) through its role as a characteristic scale in the dimensionless argument of the scaling function \mathcal{G} . Finally, and most importantly, the scaling function \mathcal{G} is not present. Hence, to derive Eq. (2) from Eq. (1), one has to make assumptions about the behaviour of \mathcal{G} in the scaling region $s_0 \ll s \ll s_c$. If the scaling function is “almost constant”, then $\tau = \tilde{\tau}$. However, if the scaling function is a power law itself, then $\tau \neq \tilde{\tau}$. To illustrate that, let us define a *cutoff function* $\tilde{\mathcal{G}}$, that, by definition, converges to a non-vanishing value for small arguments and decays rapidly for large arguments, and assume that the scaling function is of the form $\mathcal{G}(x) = x^\alpha \tilde{\mathcal{G}}(x)$. Since the cutoff function $\tilde{\mathcal{G}}$ is constant and non-zero in the scaling region, the PDF behaves like $as^{-\tau} (s/s_c)^\alpha \tilde{\mathcal{G}}(0)$ for $s_0 \ll s \ll s_c$, so that the scaling exponent τ in Eq. (1) is related to the apparent exponent $\tilde{\tau}$ in Eq. (2) via $\tau - \alpha = \tilde{\tau}$.

One of the key problems found in the literature is the identification of the scaling exponent τ with the apparent exponent $\tilde{\tau}$; however, they differ if the scaling function has a power-law dependence on its argument. First, we illustrate the consequences of such a scaling function further, concluding that if the apparent exponent $\tilde{\tau}$ in Eq. (2) is less than unity, then the scaling exponent τ in Eq. (1) is exactly 1. Next, this result is derived in general by putting the formalism, Eqs. (1) and (2), on a sound mathematical basis, generating further general properties in the process such as $\tau \geq 1$, $\tau \geq \tilde{\tau}$ and $\tilde{\tau} \geq 1 \Rightarrow \tau = \tilde{\tau}$. Finally, we review examples from the literature that illustrate the formalism developed, including some that suffer from a confusion of τ and $\tilde{\tau}$.

1.1 Example

To demonstrate the first result explicitly, we assume that the PDF for the observable $s \in [s_0, \infty[$ has the form

$$P(s; s_c) = \tilde{a}s^{-\tilde{\tau}}\tilde{\mathcal{G}}(s/s_c) \quad \text{for } s \geq s_0 = 1, \quad (3)$$

see Fig. 1(a). Motivated by a previous study [7], in this example, the cutoff function is assumed to be $\tilde{\mathcal{G}}(x) = (1-x)^{\tilde{\tau}}\theta(1-x)$ with θ being the Heaviside step function. (A simpler, but more artificial example would be to consider a cutoff function identical to a pure Heaviside step function.) Along the lines of the considerations above, this cutoff function converges to unity for small arguments and vanishes for all $x > 1$. The apparent exponent, the slope of $P(s; s_c)$ in the scaling region as shown in Fig. 1(a), therefore is the exponent $\tilde{\tau}$ in Eq. (3). A naïve comparison of Eq. (1) and Eq. (3) suggests that the scaling exponent τ is the slope of a straight-line fit to $P(s; s_c)$ vs. s in a double logarithmic plot, but this is incorrect if $\tilde{\tau} < 1$, as we will see in the following.

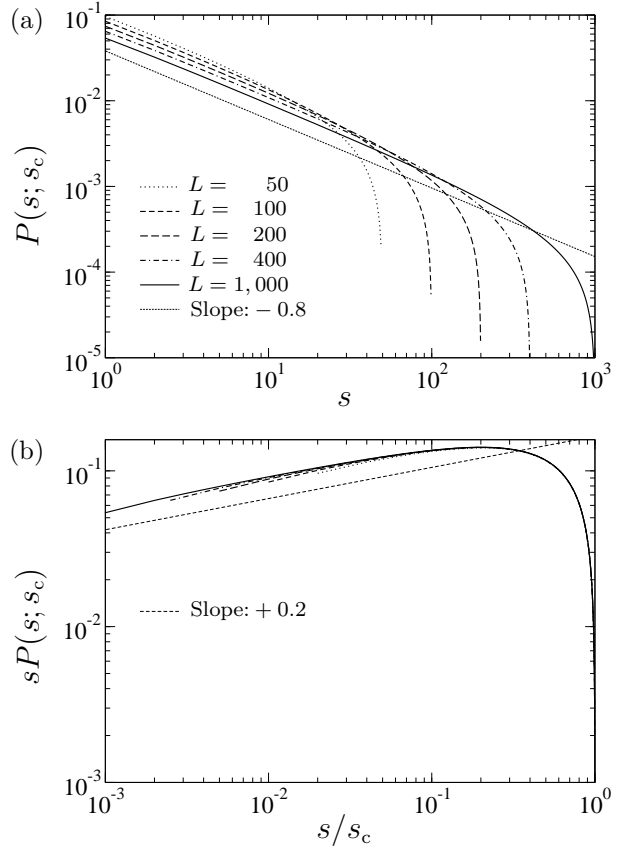


Fig. 1. (a) The PDF $P(s; s_c)$ vs. s with an apparent exponent $\tilde{\tau} = 0.8$ (indicated by a straight line with slope -0.8). Since $\tilde{\tau} < 1$, the prefactor \tilde{a} decreases with increasing cutoff, shifting the graphs downward the ordinate, see Eq. (4b). (b) Data collapse. The transformed PDF $sP(s; s_c)$ vs. the rescaled observable s/s_c . For $s, s_c \gg s_0$ the curves collapse onto the graph for the scaling function $\mathcal{G}(x) = x^{1-\tilde{\tau}}\tilde{\mathcal{G}}(x)$ where the cutoff function $\tilde{\mathcal{G}}(x) = (1-x)^{\tilde{\tau}}\theta(1-x)$, see Eq. (3). For small arguments the scaling function behaves like a power law $x^{0.2}$, as indicated by a straight line with slope $+0.2$. Example taken from Ref. [7].

In this particular example (3), in the limit of large s_c the scaling form of the PDF is expected to account for all s , so that in this limit

$$1 = \int_{s_0=1}^{\infty} ds P(s; s_c) = \begin{cases} \tilde{a}^{-1} & \text{for } \tilde{\tau} > 1, \\ \tilde{a}s_c^{1-\tilde{\tau}}\Gamma(1-\tilde{\tau})\Gamma(1+\tilde{\tau}) & \text{for } \tilde{\tau} < 1. \end{cases} \quad (4)$$

For simplicity, we ignore the case $\tilde{\tau} = 1$ which contains logarithmic corrections. Equation (4) determines \tilde{a} , which for $\tilde{\tau} < 1$ cannot be a constant but has to vanish like $s_c^{\tilde{\tau}-1}$ in the limit of large s_c . Substituting \tilde{a} into Eq. (3) and rearranging, the PDF reads

$$P(s; s_c) = \begin{cases} (\tilde{\tau}-1)s^{-\tilde{\tau}}\tilde{\mathcal{G}}(s/s_c) & \text{for } \tilde{\tau} > 1, \\ \frac{1}{\Gamma(1-\tilde{\tau})\Gamma(1+\tilde{\tau})}s^{-1}(s/s_c)^{1-\tilde{\tau}}\tilde{\mathcal{G}}(s/s_c) & \text{for } \tilde{\tau} < 1. \end{cases} \quad (5)$$

By comparing Eq. (5) with the FSS ansatz (1), we see that when the apparent exponent $\tilde{\tau} > 1$, the metric factor $a = (\tilde{\tau} - 1)$, the scaling exponent $\tau = \tilde{\tau}$, and the scaling function $\mathcal{G}(x) = \tilde{\mathcal{G}}(x)$. However, if the measured apparent exponent $\tilde{\tau} < 1$, the metric factor $a = (\Gamma(1 - \tilde{\tau})\Gamma(1 + \tilde{\tau}))^{-1}$, the scaling exponent $\tau = 1$, and the scaling function $\mathcal{G}(x) = x^{1-\tilde{\tau}}\tilde{\mathcal{G}}(x)$. For the FSS ansatz (1), it is a necessity that the metric factor is asymptotically independent of s_c . Dropping this constraint would render the *definition* of the scaling exponent τ meaningless: One may rewrite $as^{-\tau}\mathcal{G}(s/s_c) = as_c^\alpha s^{-(\tau+\alpha)}(s/s_c)^\alpha \mathcal{G}(s/s_c)$ for any α . The factor $(s/s_c)^\alpha$ could be absorbed into the scaling function. Were it allowed to absorb the power s_c^α in the metric factor, it would be impossible to differentiate between τ and $\tau + \alpha$.

2 Derivation

Now we present a more general and rigorous derivation of the result illustrated above.

2.1 Definitions and notation

The FSS ansatz (1) can be recast more formally as follows: For all $\epsilon > 0$ there exists a constant lower cutoff s_0 and an \mathcal{S} so that for fixed $n \geq 0$, the relative error of the n th moment from the approximate PDF of the observable $s \in [0, \infty[$

$$P_{\text{app}}(s; s_c) = \begin{cases} as^{-\tau}\mathcal{G}(s/s_c) & \text{for } s \geq s_0, \\ f(s, s_c) & \text{otherwise} \end{cases} \quad (6a)$$

$$(6b)$$

compared to the exact n th moment of the PDF approximated by $P_{\text{app}}(s; s_c)$ is less than or equal to ϵ for all $s_c > \mathcal{S}$. Here, we have introduced $f(s, s_c)$, the non-universal part of the PDF. Both functions, $f(s, s_c)$ and $\mathcal{G}(x)$, are non-negative, and the latter is assumed to be non-zero at least in a finite range of x . We assume $s_0 > 0$, but the case $s_0 = 0$ is briefly discussed as well.

The exact n th moment is given by

$$\langle s^n \rangle = f_n(s_0, s_c) + as_c^{1+n-\tau}g_n(s_0/s_c) + \mathcal{O}(\epsilon) \quad (7)$$

where

$$f_n(s_0, s_c) = \int_0^{s_0} ds s^n f(s, s_c), \quad (8a)$$

$$g_n(s_0/s_c) = \int_{s_0/s_c}^{\infty} dx x^{n-\tau} \mathcal{G}(x). \quad (8b)$$

Both f_n and g_n are non-negative and g_n is a monotonically non-increasing function of s_0/s_c . It is worth noting that the range over which the scaling region is probed widens as s_0/s_c approaches 0.

2.2 Scaling exponent $\tau \geq 1$

Since g_0 does not decrease with s_c , the normalisation condition

$$1 = \langle s^0 \rangle = f_0(s_0, s_c) + as_c^{1-\tau}g_0(s_0/s_c) + \mathcal{O}(\epsilon) \quad (9)$$

implies $\tau \geq 1$, otherwise $as_c^{1-\tau}g_0(s_0/s_c)$ would diverge as $s_c \rightarrow \infty$. *Therefore, the scaling exponent $\tau \geq 1$.*

2.3 If \mathcal{G} is continuous, the limit $\mathcal{G}_0 = \lim_{x \rightarrow 0} \mathcal{G}(x)$ exists and is finite

The behaviour of the scaling function \mathcal{G} in the limit of small arguments s/s_c enters into the scaling of the moments in the asymptotic limit of large upper cutoff (the thermodynamic limit). We will assume that $\mathcal{G}(x)$ is continuous within a finite range $]0, q]$ with $q > 0$ and that the integrals in Eq. (8) exist. Without these (reasonable) assumptions, the discussion of the behaviour of the integral g_n becomes too complicated. The continuity implies further that $\mathcal{G}(x)$ cannot diverge with small arguments: If it did, $g_0(s_0/s_c)$ diverges faster than $s_c^{\tau-1}$ for $\tau > 1$ and at least logarithmically fast for $\tau = 1$, contradicting the normalisation condition Eq. (9). *Therefore, if $\mathcal{G}(x)$ is continuous, it cannot be divergent in small arguments. Together with the continuity, the limit $\mathcal{G}_0 \equiv \lim_{x \rightarrow 0} \mathcal{G}(x)$ therefore exists and is finite.*

2.4 If $\tau = 1$ then $\mathcal{G}_0 = 0$

We consider the case $\mathcal{G}_0 > 0$: If $\mathcal{G}(x)$ is non-zero and bounded from above for small $x \geq 0$, the integral $g_n(s_0/s_c)$ is logarithmically divergent in s_c when $1 + n - \tau = 0$, has leading order $s_c^{-(1+n-\tau)}$ when $1 + n - \tau < 0$ and converges to a non-zero value with increasing s_c when $1 + n - \tau > 0$. The normalisation condition Eq. (9) implies $as_c^{1-\tau}g_0(s_0/s_c) \leq 1$, so that $\mathcal{G}_0 > 0$ violates the normalisation condition if $\tau \leq 1$. For $\tau > 1$, the divergence $s_c^{-(1-\tau)}$ of g_0 is cancelled by the prefactor $s_c^{1-\tau}$. *Therefore, if $\mathcal{G}_0 > 0$ then $\tau > 1$. Negating this result yields: If $\tau = 1$ then $\mathcal{G}_0 = 0$.*

2.5 If the lower cutoff $s_0 = 0$ then $\tau = 1$ and $\mathcal{G}_0 = 0$

The result above can be amended by the observation that if the lower cutoff $s_0 = 0$ then all non-universal f_n vanish and the normalisation condition Eq. (9) reduces to $1 = as_c^{1-\tau}g_0(0) + \mathcal{O}(\epsilon)$, which can hold only if $\tau = 1$. Moreover, $g_0(0) = \int_0^\infty dx x^{-1}\mathcal{G}(x)$ can converge for continuous $\mathcal{G}(x)$ only if $\mathcal{G}_0 = 0$, consistent with our findings in Section 2.4 *Therefore, if the lower cutoff vanishes, $s_0 = 0$, then $\tau = 1$ and $\mathcal{G}_0 = 0$.* Generally, the converse does not hold, that is, $\mathcal{G}_0 = 0$ does not imply $s_0 = 0$. However, $\mathcal{G}_0 > 0$ or $\tau > 1$ implies $s_0 > 0$.

2.6 If $\mathcal{G}_0 = 0$ then $\tau = 1$

We consider the case $\mathcal{G}_0 = 0$: Here, we focus on g_0 . If $\tau > 1$, then $g_0(s_0/s_c)$ diverges slower than $(s_0/s_c)^{1-\tau}$ in decreasing s_0/s_c . This can be proven by assuming the opposite, implying that there exists a $c > 0$ so that $\int_{s_0/s_c}^{\infty} dx x^{-\tau} \mathcal{G}(x) \geq c(s_0/s_c)^{1-\tau}$ for all s_0/s_c below a threshold determined by c . The right hand side of this relation can be written as $(\tau - 1) \int_{s_0/s_c}^{\infty} dx x^{-\tau}$. Independent of its prefactor, for sufficiently small x the integrand represents an upper bound, given that $\mathcal{G}(x)$ is less than any δ for sufficiently small x . It is easily shown that the above relation therefore eventually breaks down, implying that the product $as_c^{1-\tau} g_0(s_0/s_c)$, the fraction of events within the universal scaling part of the PDF $P(s; s_c)$, vanishes in increasing s_c if $\tau > 1$. No such argument exists if $\tau = 1$. Mathematically, a vanishing fraction of such “universal events” constraints neither $\mathcal{G}(x)$ nor τ . However, it makes sense to demand some weight in the scaling part of the PDF on physical grounds. *Therefore, if \mathcal{G}_0 vanishes and the fraction of events in the scaling part of the PDF is finite in the thermodynamic limit, then the scaling exponent $\tau = 1$. If \mathcal{G}_0 vanishes and the scaling exponent $\tau > 1$, then the fraction of events in the scaling part of the PDF vanishes in the thermodynamic limit.* Hence, we are faced with the almost paradoxical situation that all moments with $n > \tau - 1$ are determined by the scaling part of the PDF, the measure of which, however, vanishes in thermodynamic limit. In this case, the normalisation condition Eq. (9) entails that $f_0(s_0, s_c)$ converges to 1 in the limit of large s_c , that is, the lower cutoff s_0 cannot vanish. Finally, we note that, if $\mathcal{G}_0 = 0$, then $g_n(s_0/s_c)$ cannot diverge faster for $s_0/s_c \rightarrow 0$ than in the case $\mathcal{G}_0 > 0$ since $\mathcal{G}(x)$ does not diverge for small arguments. *Therefore, if $\mathcal{G}_0 = 0$ then $g_n(s_0/s_c)$ converges for all $n > \tau - 1$.*

2.7 If $\mathcal{G}_0 > 0$ then $\tau = \tilde{\tau}$

The apparent exponent $\tilde{\tau}$ is formally defined as

$$\lim_{s \rightarrow \infty} \lim_{s_c \rightarrow \infty} \frac{P(\lambda s; s_c)}{P(s; s_c)} \lambda^{\tilde{\tau}} = 1 \quad \text{for all } \lambda > 0, \quad (10)$$

if the limits and such an exponent exist, which is, for example, not the case if $\mathcal{G}(x)$ converges to 0 faster than any power law. The limit $s \rightarrow \infty$ guarantees that the limit $s_c \rightarrow \infty$ is taken in the scaling region, so that $P(s; s_c)$ can be replaced by the scaling part Eq. (6a):

$$\lim_{s \rightarrow \infty} \lim_{s_c \rightarrow \infty} \frac{\mathcal{G}(\lambda s/s_c)}{\mathcal{G}(s/s_c)} \lambda^{\tilde{\tau}-\tau} = 1 \quad \text{for all } \lambda > 0. \quad (11)$$

If $\mathcal{G}_0 > 0$, the fraction in Eq. (11) converges to 1 and for the left hand side to be equal to unity for all λ , the exponent $\tilde{\tau} - \tau$ vanishes. *Therefore, if $\mathcal{G}_0 > 0$ then $\tau = \tilde{\tau}$ or, equivalently, $\tau \neq \tilde{\tau} \Rightarrow \mathcal{G}_0 = 0$.*

2.8 Generally $\tau \geq \tilde{\tau}$, if $\tilde{\tau} < 1$ then $\tau = 1$, and if $\tilde{\tau} \geq 1$ then $\tau = \tilde{\tau}$

Equation (11) implies that $\mathcal{G}(\lambda s/s_c)$ scales like $\lambda^{\tau-\tilde{\tau}}$ and therefore diverges in small arguments if $\tau < \tilde{\tau}$. However, as noted previously, \mathcal{G} cannot diverge for small arguments. *Therefore, we have the general result $\tau \geq \tilde{\tau}$.*

Moreover, the cutoff function $\tilde{\mathcal{G}}(x) = x^{\tilde{\tau}-\tau} \mathcal{G}(x)$ can then be expected to converge to a (non-vanishing) constant as $x \rightarrow 0$, and imposing that, one can define uniquely the apparent exponent $\tilde{\tau}$ through

$$P(s; s_c) = \tilde{a}(s_c) s^{-\tilde{\tau}} \tilde{\mathcal{G}}(s/s_c) \quad \text{for } s \geq s_0. \quad (12)$$

Using the above two arguments, the scaling exponent τ and the apparent exponent $\tilde{\tau}$ can be related as follows: If $\tilde{\tau} < 1$ then $\tau \neq \tilde{\tau}$ since $\tau \geq 1$, implying $\mathcal{G}_0 = 0$. As shown above, $\mathcal{G}_0 = 0$ implies $\tau = 1$ assuming that the scaling part of the PDF is non-empty. *Therefore, if $\tilde{\tau} < 1$ then $\tau = 1$.*

Now, assume to the contrary that $\tau \neq \tilde{\tau} \geq 1$. Hence $\mathcal{G}_0 = 0$ and again, assuming a non-empty scaling part of the PDF, $\tau = 1$. If $\tilde{\tau} = 1$ that clashes with $\tau \neq \tilde{\tau}$, if $\tilde{\tau} > 1$ that clashes with $\tau \geq \tilde{\tau}$. *Therefore, if $\tilde{\tau} \geq 1$ then $\tau = \tilde{\tau}$, unless the total measure of the scaling part of the PDF vanishes asymptotically.*

3 Discussion

In this section, we first discuss some implications of the above derivations and then present some examples from the literature, where a more thorough scaling analysis leads to revised results.

In many models of equilibrium [8] as well as non-equilibrium [9] statistical mechanics, within the FSS regime, the second moment $\langle s^2 \rangle$ of the total order parameter¹ scales like its first moment squared $\langle s \rangle^2$, which is usually linked to so-called hyper-scaling. This implies that their ratio converges in the thermodynamic limit, that is,

$$\lim_{s_c \rightarrow \infty} \frac{\langle s^2 \rangle}{\langle s \rangle^2} = r, \quad (13)$$

where r is a finite number. We now show that such finite r implies $\tau = 1$, assuming that Eq. (1) holds in these models, by showing that if $\tau > 1$, then the ratio diverges asymptotically. To invoke the argument “if $\tau > 1$ then $\mathcal{G}_0 > 0$ ”, we need to impose that a non-vanishing fraction of the total order parameter distribution is governed by finite-size scaling. In equilibrium statistical mechanics, that means that the singular part of the partition sum does not vanish asymptotically. Moreover, we assume that both the first and the second moment diverge in the thermodynamic limit. This is important, because r may be

¹ We distinguish between the total order parameter s and the order parameter density s/L^d . The analysis is identical for the order parameter density, however, the upper cutoff then vanishes asymptotically which would require a revision of previous arguments.

finite even for $\mathcal{G}_0 > 0$ simply by convergence of the first two moments if $\tau > 3$, as can be shown using Eq. (7) and the divergence of g_n as discussed in Sec.2.7. Based on the arguments presented there, if the first moment diverges, then $\tau \leq 2$. Using the same arguments, one can now show that for $2 > \tau > 1$ the ratio r diverges like $s_c^{\tau-1}$ and for $\tau = 2$ it diverges like $s_c^{\tau-1}$ with logarithmic corrections. Thus asymptotically finite r is incompatible with $\tau > 1$, if the first moment diverges and the fraction of the total order parameter distribution that is governed by the scaling part is finite. The only alternative is $\tau = 1$, in which case $\mathcal{G}_0 = 0$ and therefore all g_n are finite for all n and $r = a^{-1}g_2/g_1^2$ from Eq. (7). The case $\tau = 1$ therefore is anything but special, describing most of standard critical phenomena. It is worth noting that $\tau = 1$ implies a dimensionless metric factor a .

We illustrate the observation above by considering the finite-size scaling of the cluster number density ² in percolation [2] (i.e., the number of clusters of size s per site in a system of linear size L)

$$n(s; s_c^{\text{perc}}) = a_{\text{perc}} s^{-\tau^{\text{perc}}} \mathcal{G}^{\text{perc}}(s/s_c^{\text{perc}}), \quad (14)$$

where $s_c^{\text{perc}} = b^{\text{perc}} L^D$, $D = 1/(\nu\sigma)$ with ν and σ being the critical exponents describing the divergence of the correlation length and the characteristic cluster size, respectively, and b^{perc} being a metric factor. The exponent τ^{perc} is different from 1, known to be exactly 187/91 in two-dimensional percolation. However, $n(s; s_c^{\text{perc}})$ is not the distribution of the total order parameter, which is the number of occupied sites in the largest cluster. Multiplying $n(s; s_c^{\text{perc}})$ by L^d , one obtains the average number of clusters of size s in a system with L^d sites. The probability of the largest cluster being of a particular size must be contained in this distribution: If $n(s; s_c^{\text{perc}})$ was the site-normalised density of the largest cluster only, then $L^d n(s; s_c^{\text{perc}})$ was the probability that the largest cluster has size s . If $\tilde{n}(s; s_c^{\text{perc}})$ is the site-normalised histogram of all clusters excluding the largest, one could therefore write $L^d n(s; s_c^{\text{perc}}) = L^d \tilde{n}(s; s_c^{\text{perc}}) + P(s; s_c^{\text{perc}})$ with $P(s; s_c^{\text{perc}})$ being the distribution of the total order parameter. One might speculate that $P(s; s_c^{\text{perc}})$ is responsible for the characteristic bump in the histogram $n(s; s_c^{\text{perc}})$. Assuming that $P^{\text{perc}}(s; s_c^{\text{perc}})$ follows the same scaling as $L^d n(s; s_c^{\text{perc}})$ itself, one has

$$P^{\text{perc}}(s; s_c^{\text{perc}}) = a'_{\text{perc}} L^d s^{-\tau^{\text{perc}}} \tilde{\mathcal{G}}^{\text{perc}}(s/s_c^{\text{perc}}), \quad (15)$$

with $s_c^{\text{perc}} = b^{\text{perc}} L^D$. Absorbing L^d into a redefinition of the scaling function $\mathcal{G}^{\text{perc}}(x) = x^{-d/D} \tilde{\mathcal{G}}^{\text{perc}}(x)$ the distribution of the order parameter scales like

$$P^{\text{perc}}(s; s_c^{\text{perc}}) = a'_{\text{perc}} s^{\frac{d}{D} - \tau^{\text{perc}}} \mathcal{G}^{\text{perc}}(s/s_c^{\text{perc}}), \quad (16)$$

where the exponent of s turns out to be $d/D - \tau^{\text{perc}} = -1$ by standard scaling laws [2], indeed as expected from the previous analysis.

² In the following, we consider finite systems at occupation probability p tuned to the critical value p_c , so that we do not have to exclude from the cluster number density the largest, infinite or percolating cluster(s).

3.1 Examples

Finally, we want to present some examples from the literature, where a misunderstanding of the theory of scaling has lead to some confusion or even erroneous conclusions. In [10] the cluster number density in steady state two-phase flow in porous media is investigated. As discussed above, in percolation the cluster number density $n(s)$ is the site normalised density of clusters of size s , so that $sn(s)$ is the probability that a randomly chosen site belongs to a cluster of size s . The histogram is expected to follow scaling, $n(s) = a^{\text{perc}} s^{\tau^{\text{perc}}} \mathcal{G}^{\text{perc}}(s/s_c^{\text{perc}})$, see Eq. (14). The sum $p = \sum_s sn(s)$ is the fraction of sites belonging to any cluster, i.e. the fraction of occupied sites. This sum is bounded and non-zero in non-trivial cases even if the largest clusters are excluded from $n(s)$, so that the results for $P(s; s_c)$ derived in this paper, translate into corresponding results for $sn(s)$. Performing the sum suggests $p \propto (s_c^{\text{perc}})^{2-\tau^{\text{perc}}}$, which means that p vanishes asymptotically for $\tau^{\text{perc}} > 2$. This, however, is a misunderstanding of the theory of scaling, because it ignores the existence of a finite lower cutoff s_0^{perc} . In fact, $\tau^{\text{perc}} \geq 2$, corresponding to $\tau \geq 1$ as derived earlier, is a necessary condition for p to be finite. The numerical finding of $\tau^{\text{perc}} \approx 1.92$ in [10] for a cluster number density therefore is theoretically impossible – in fact it is based on a straight line fit in a double logarithmic plot and therefore it is $\tilde{\tau}$ and not τ that is estimated.

The term self-organised criticality refers to the tendency of slowly driven non-equilibrium systems with many degrees of freedom to spontaneously reach a critical state where the PDF of relaxational event sizes obey FSS. There are only a few controlled laboratory experiments on scale-invariance in slowly-driven non-equilibrium systems. In one such experiment, the avalanche-size PDF was measured in slowly driven piles of rice [11]. For piles with elongated grains, the avalanche-size PDF obeys FSS. However, it was claimed that a system with more spherical grains is non-critical, since the PDF of the avalanche size s in a system of size L is consistent with $P(s; L) \propto L^{-1} \tilde{\mathcal{G}}(s/L)$ where the function $\tilde{\mathcal{G}}(x) \propto \exp[-(x/x^*)^\gamma]$ with $\gamma \approx 0.43$ and $x^* \approx 0.45$ was identified as a scaling function rather than a cutoff function. This is a misunderstanding of the theory of scaling. The PDF can be recast into the FSS form $P(s; L) \propto s^{-1} \mathcal{G}(s/L)$ where the scaling function $\mathcal{G}(x) = x \tilde{\mathcal{G}}(x)$ increases linearly with x for small arguments, decaying (stretched) exponentially fast for large arguments. Hence, also for the piles with more spherical grains does the avalanche-size PDF obey FSS but with a scaling exponent $\tau = 1$ and $\mathcal{G}_0 = 0$. Therefore, the occurrence of self-organised criticality does not depend on the details of the system as wrongly claimed in [11].

Complex networks have attracted intense interest from the statistical physics community in particular because they often exhibit scale invariance. It has been suggested that so-called duplication-divergence-mutation models, where nodes correspond to proteins and links correspond to pairwise interactions between proteins, are good candidates to describe the evolution and large-scale topological fea-

tures of real protein-protein interaction networks. A recent study [12] derives asymptotic properties for such models and it is reported that the degree PDF for a node to have degree k at time t is $P(k; t) \propto t^{-1}(k/t)^{k_{\min}-1}\tilde{\mathcal{G}}(k/t)$ where k_{\min} is the lower non-zero degree in the initial configuration, $\tilde{\mathcal{G}}$ a cutoff function, and it is claimed that the scaling exponent is $k_{\min} - 1$ [12]. However, this is a misunderstanding of the theory of scaling. A careful analysis reveals that the PDF can be recast into the FSS form $P(k; t) = ak^{-1}\mathcal{G}(k/t)$ where the metric factor $a = k_{\min}(1 + k_{\min})$ and the scaling function $\mathcal{G}(x) = x^{k_{\min}}(1-x)\theta(1-x)$ increases like a power law with exponent k_{\min} with x for small arguments, decaying fast for large arguments. Hence, the FSS scaling exponent $\tau = 1$, independent of k_{\min} and $\mathcal{G}_0 = 0$.

Earthquake statistics has been a prominent subject of statistical analysis by means of scaling arguments (*e.g.* [13]). In [14], the distribution $P_{m,L}(\Delta r)$ of distances of successive earthquakes with a magnitude greater than m within a cell of linear L is considered. The analysis shows that $P_{m,L}(\Delta r) = L^{-1}f(\Delta r/L)$, where $f(x) \propto x^{-0.6}$. As that implies $P_{m,L}(\Delta r) \propto \Delta r^{-0.6}$, the distribution seems to be asymptotically non-normalisable. This conclusion, however, is a misunderstanding of the theory of scaling and a confusion of $\tilde{\tau}$ and τ : The PDF $P_{m,L}(\Delta r)$ was non-normalisable only if $\tau < 1$, but in fact $\tau = 1$, which can be seen by rewriting the scaling of the PDF as $P_{m,L}(\Delta r) = \Delta r^{-1}\mathcal{G}(\Delta r/L)$ with $\mathcal{G}(x) = xf(x)$. The scaling function $\mathcal{G}(x)$ converges to 0 for small arguments like $x^{1-0.6}$, as expected from the discussion above. In the introduction we showed that $\mathcal{G}(x) \propto x^\alpha$ implies $\tilde{\tau} = \tau - \alpha$, so that in fact $\tilde{\tau} = 0.6$ in the present case.

In summary, given the probability density function $P(s; s_c)$ obeys the FSS ansatz (1), if the apparent exponent $\tilde{\tau}$ of $P(s; s_c)$ in Eq. (2) is less than unity, then the scaling exponent τ is unity and the scaling function vanishes for small arguments, $\lim_{x \rightarrow 0} \mathcal{G}(x) = 0$. If $\tilde{\tau} \geq 1$, then the scaling exponent equals the apparent exponent, $\tau = \tilde{\tau}$. If $\tau > 1$, then the scaling function approaches a non-zero value for small arguments, $\lim_{x \rightarrow 0} \mathcal{G}(x) > 0$.

GP thanks Mervyn Freeman, Nicholas Watkins and Max Werner for useful discussions.

References

1. H.E. Stanley, *Introduction to Phase Transitions and Critical Phenomena* (Oxford University Press, 1971)
2. D. Stauffer, A. Aharony, *Introduction to Percolation Theory* (Taylor & Francis London, 2nd edition, 1994)
3. H.J. Jensen, *Self-Organised Criticality* (Cambridge University Press, 1998)
4. F.J. Wegner, Phys. Rev. B **5**, 4529 (1972)
5. V. Privman, P.C. Hohenberg, A. Aharony, in *Phase Transitions and Critical Phenomena*, edited by C. Domb, J.L. Lebowitz (Academic, New York, 1991), Vol. 14, chap. 1, pp. 1–134
6. R. Lübeck, R.D. Willmann, Nucl. Phys. B **718**, 341 (2005)
7. N. Farid, K. Christensen, New J. Phys. B **8**, 212 (2006)
8. P. Pfeuty, G. Toulouse, *Introduction to the Renormalization Group and to Critical Phenomena* (John Wiley & Sons, Chichester, 1977)
9. J. Marro, R. Dickman, *Nonequilibrium Phase Transitions in Lattice Models* (Cambridge University Press, New York, 1999)
10. T. Ramstad, A. Hansen, Phys. Rev. E **73**, 026306 (2006)
11. V. Frette, K. Christensen, A. Mølthe-Sørensen, J. Feder, T. Jøssang, P. Meakin, Nature **379**, 49 (1996)
12. A. Raval, Phys. Rev. E **68**, 066119 (2003)
13. P. Bak, K. Christensen, L. Danon, T. Scanlon, Phys. Rev. Lett. **88**, 178501 (2002)
14. J. Davidsen, M. Paczuski, Phys. Rev. Lett. **94**, 048501 (2005)